

Rational Design and In-Silico Evaluation of a Multi-Epitope Vaccine Targeting Human Glypican-3 (GPC3) for Hepatocellular Carcinoma Immunotherapy

Rawaa Najim Alkhamessi¹, Fatima H. B. Muarid², Aliaa M.M. Khalifa³, Salma Elnour Rahma⁴, Mohammed Ibrahim Saeed⁴, Abdelgadir Elamin Eltom⁴, Marwan Ismail⁴, Ahmed L. Osman⁴, Sara Mohammed Ali⁴

¹College of Science, Mustansiriyah University, Baghdad, Iraq

²Department of Entomology and Invertebrates, University of Baghdad (Iraq Natural History Research Center and Museum, Baghdad, Iraq)

³Department of chemistry, College of Art and science, AL-Marj, Benghazi University, AL-Marj, Libya

⁴Gulf Medical University, Jurf, Ajman, UAE

Emails:

Rawaa Najim Alkhamessi: dr.rawaanajim@uomustansiriyah.edu.iq

Fatima H. B. Muarid: fatima.h@nhm.uobaghdad.edu.iq

Aliaa M.M.Khalifa: Oliakalifa.mrr1980@gmail.com

Salma Elnour Rahma: dr.salmaelnour@gmu.ac.ae

Mohammed Ibrahim Saeed: moibsaheed@yahoo.com

Abdelgadir Elamin Eltom: gadoora1977@live.com

Marwan Ismail: marwan@gmu.ac.ae

Ahmed L. Osman: dr.ahmed@gmu.ac.ae

Sara Mohammed Ali: dr.sara@gmu.ac.ae

ABSTRACT

Hepatocellular carcinoma (HCC) is among the most common and lethal cancers in the world, with more than 900,000 new cases diagnosed every year and remains one of the leading death factors of cancer. The lack of efficacy with currently available treatments has created a need for new immunotherapeutic approaches. The cancer-testis antigen, glypican-3 (GPC3), is an oncofetal protein overexpressed in > 70% of HCC but not expressed in normal adult liver and was chosen as the objective for the rational vaccine design. B-cell, cytotoxic T-lymphocyte (CTL), and helper T-lymphocyte (HTL) epitopes were predicted with an integrated immunoinformatic pipeline, filtered for antigenicity, non-allergenicity, and non-toxicity from available databases, constructed as a stable multi-epitope structure along with appropriate linkers and adjuvants. The stability and accuracy of the construct were verified by structural modeling and validation, and its binding affinity to TLR4 was demonstrated by docking with strong hydrogen bonding interactions in combination with hydrophobic contacts. Molecular dynamics (500 ns) simulations showed conformational stability, chain compaction of the vaccine as well as thermodynamically favorable interactions. CAI for codon optimized ampC was 0.87 and optimal GC content for expression in *E. coli*. Simulation of the immune response predicted strong activations of both B and T cells, antibody production and memory. The results above clearly indicate that the GPC3-derived multi-epitope vaccine may be a good candidate for HCC immunotherapy and it deserves experimental verification.

KEYWORDS: Hepatocellular carcinoma, Glypican-3, Multi-epitope vaccine, Immunoinformatics, Molecular docking, Molecular dynamics simulation, Codon optimization, Immune simulation.

How to Cite: Rawaa Najim Alkhamessi, Fatima H. B. Muarid, Aliaa M.M.Khalifa, Salma Elnour Rahma, Mohammed Ibrahim Saeed, Abdelgadir Elamin Eltom, Marwan Ismail, Ahmed L. Osman, Sara Mohammed Ali., (2025) Rational Design and In-Silico Evaluation of a Multi-Epitope Vaccine Targeting Human Glypican-3 (GPC3) for Hepatocellular Carcinoma Immunotherapy, Vascular and Endovascular Review, Vol.8, No.10s, 203-216.

INTRODUCTION

Hepatocellular carcinoma (HCC) is the most prevalent type of primary liver cancer and a significant global health problem, it being the third cause of cancer-related death worldwide. GLOBOCAN 2020 statistics on HCC incidence are in excess of 900,000 new cases a year and such high incidence is disproportionately occurs in Asia and Sub-Saharan Africa due to endemic chronic hepatitis B and C infection alcohol consumption and metabolic syndromes [1]. Although surgical resection, radiofrequency ablation (RFA) and cytotoxic systemic therapy using sorafenib or immunotherapy have extended the median survival of patients, there continues to be poor long-term prognosis with a five-year overall survival rate of below 20% due to late presentation at an advanced stage, frequent recurrence and limited treatment regimens available. This highlights the urgent demand for novel or more effective treatment modalities against HCC to improve patient prognosis [2].

One of the potential molecular targets for HCC is Glypican-3 (GPC3), which has been recognized as a very promising biomarker and therapeutic target. Immunohistochemistry GPC3 is an oncofetal protein, a member of glypican family of heparan sulfate proteoglycan, which has important role in the regulation of cell proliferation and growth [3]. Although its normal expression is nonexistent or minimal in adult liver, GPC3 overexpression occurs aberrantly in 70–80% of HCC cases, rendering it a useful diagnostic marker and potentially a therapeutic target [4]. Critical in terms of clinical significance in liver cancer, its overexpression is correlated with an unfavorable outcome, more aggressive tumor behavior and immune escape; therefore, GPC3

may be employed as a novel diagnostic and therapeutic target for early diagnosis/screening and targeted therapy for liver tumors. Multiple experimental strategies, such as monoclonal antibodies (mAbs) and CAR-T cells targeting GPC3 have been tested with limited success due to immunogenicity, off-target toxicity, and the high cost of production [5].

This consideration highlights epitope-based vaccines as a plausible, rational, and less resource-intensive option for HCC immunotherapy. Multi-epitope vaccines utilize immunodominant B-cell, CTL, and HTL epitopes identified and assembled into a single construct to elicit a more expansive and durable immune response [6]. In contrast to the drawbacks associated with traditional vaccines, multi-epitope constructs are generally more specific, less allergenic, and associated with minimal adverse effects. Furthermore, advanced computational tools enable rapid design, optimization, and in silico testing of these constructs before experimental validation, which ultimately, accelerates and reduces costs in vaccine generation. Immunoinformatics tools have revolutionized vaccine design by enabling high-precision identification of antigenic epitopes, prediction of induced immune responses, structure modelling, and finally, molecular docking to assess receptor-ligand interactional stability [5]. Computational methods such as molecular dynamics and binding free energy calculations make it possible to assess the dynamical behavior and potential immunogenicity of proposed designs at physiological temperatures. These approaches have significantly enhanced the accuracy of vaccine design by minimizing chances of failure in subsequent preclinical and clinical testing. In the recent vaccine development era, several in silico developed vaccines have shown promising results against infections and cancers, emphasizing the critical and translational role of these computational pipelines in next-generation vaccine development [7].

The current study thus implemented an integrative rational immunoinformatics software solution to model a multi-epitope vaccine targeting human GPC3 for practical HCC immunotherapy. The selected B-cell, CTL, and HTL epitopes were rigorously screened for antigenicity, non-allergenicity, and toxicity, and further compiled to construct a sustainable vaccine model with appropriate linkers and adjuvants. Further, structural prediction, molecular docking, and molecular dynamics simulations were implemented to analyze the immune receptor and vaccine interaction and binding and evaluate the vaccine's immunogenic and physicochemical properties. This in silico approach combining multiple methods may lead to the cost effectiveness and robust vaccine model, with efficient HCC immunotherapy targets completely validating the current clinical constraints. Above all, children represent a broad in silico implementation and a prime field of usage.

METHODOLOGY

The development and testing of a multi-epitope vaccine candidate was implemented using an extensive computational immunoinformatic pipeline. The workflow was integrated with a suite of bioinformatics tools and internet-based servers to ensure adequate prediction, selection and validation of epitopes, structural and immunological characterization of the vaccine construct. All of the procedures, beginning with the rescue of the viral protein sequences and immunogenicity studies, were carefully considered to ensure the highest level of safety, stability, and immunogenicity.

Retrieval and Selection of Target Proteins

Viral proteomic sequences of the viral strains were obtained as FASTA format files in the UniProt database (<https://www.uniprot.org>). UniProt is an all-purpose source of protein sequence and function information. It enables retrieval of the protein sequence which is the first step in the development of a vaccine [8].

Selection of targeted protein

The selection criteria of the appropriate target proteins were based on the antigenicity, allergenicity and toxicity profiles. Antigenicity was assessed utilizing the VaxiJen server (<http://www.ddg-pharmfac.net/vaxijen/>), allergenicity was evaluated through the AllerTOP v2.0 platform (<https://www.ddg-pharmfac.net/AllerTOP/>), and toxicity was analyzed via the ToxinPred tool (<https://webs.iitd.edu.in/raghava/toxinpred/>) [9], [10]. Transmembrane topology and subcellular localization were ascertained using the TMHMM server to ensure that the selected proteins were accessible and suitable for epitope mapping [11].

Epitope Prediction and Filtration

The B cell, MHC class I and class II T-cell epitopes were identified by using an analysis resource, Immune Epitope Database (IEDB) (<http://tools.iedb.org/main/tcell/>). The IEDB tools that applied ANN 4.0 to the analysis of MHC class I and NN-align 2.3 to MHC class II prediction were used to identify an epitope. Epitopes that had low IC50 and high immunogenicity prediction were chosen to be further analyzed [12].

Vaccine Construct Assembly

Specific linkers were inserted to allow successful linking of the chosen epitopes in a progressive manner to enhance the right folding and to minimize steric interference. MHC class I and II epitopes are linked using SSL and AAY which had flexibility and physical separation requirements. The adjuvant that was used and is immuno stimulatory was Patent US11077184 inserted through an EAAAK peptide linkage. The lipidated peptides resemble lipoproteins of bacteria and are able to stimulate the activation of Toll-like receptor 2 (TLR2) in the enhancement of the immune system. The C-terminus was also extended by a 6xHis linked using kk linker, and a removal signal (VS) to facilitate expression and purification. Immunogenicity of the antigenic proteins, allergenicity and toxicity of the construct was tested linearly to check the safety and probably efficacy [13].

Vaccine allergenicity and antigenicity prediction:

The allergenicity of the vaccine was evaluated through the utilization of the AllerTOP v.2.0 server (<https://www.ddg-pharmfac.net/AllerTOP/>). This methodology converts protein sequences into standardized vectors of uniform length by

leveraging auto cross-covariance. The outcomes of this analysis indicate whether the protein exhibits non-allergenic or allergenic properties. The antigenicity of the vaccine was assessed utilizing the VaxiJen web server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>). This server employs a strategy that obviates the necessity for alignment, concentrating instead on the primary characteristics of amino acids, with a specified threshold [5].

Solubility prediction

The solubility of the vaccine construct was determined with the help of the Protein-Sol tool (<https://protein-sol.manchester.ac.uk/>) that predicts the solubility of the protein using the amino acid sequence. Protein-Sol applies computational involved rules to compute a prediction of likelihood that a specified protein is soluble or not at certain conditions [14].

Physiochemical properties

A computational tool referred to as ProtParam (<https://web.expasy.org/protparam/>) was employed to evaluate physicochemical properties. This tool provides comprehensive information regarding the dimensions of the vaccine's intended structure, including metrics such as number of residues, solubility, molecular weight, extinction coefficient and isoelectric point [15].

Structural and Biophysical Evaluation

The secondary structure of the vaccine construct was anticipated using the PSIPRED server (<http://bioinf.cs.ucl.ac.uk/psipred/>), which classifies sequences into alpha-helices, beta-strands, and coils [16].

Tertiary structure prediction

The tertiary structure was modeled employing the AlphaFold server 3 (<https://alphafoldserver.com/>), which predicts inter-residue orientations to generate accurate three-dimensional models [17].

Validation of protein structure

Structural validation was conducted utilizing PROCHECK (<http://services.mbi.ucla.edu/SAVES/>) and the SAVES server to produce Ramachandran plots and ERRAT scores, which evaluate stereochemical quality and model accuracy [18].

Immune Interaction Analysis

In order to assess the binding interactions between the vaccine candidate and human immune receptors, molecular docking studies were conducted utilizing either the ClusPro 2.0 (https://cluspro.org/tut_dock.php) server [19]. The docked complexes were subjected to analysis regarding interaction energies, confidence scores, and the quantity of hydrogen bonds formed. The LigPlot+ software (<https://www.ebi.ac.uk/thornton-srv/software/LigPlus/>) provided comprehensive visualizations of molecular interactions, elucidating hydrogen bond networks and binding residues that are pivotal to the vaccine-receptor interaction [20].

Molecular Dynamics Simulations

The dynamic stability and interaction pattern of the protein–protein complex was studied through molecular dynamics simulation with Desmond (Schrödinger LLC). The docked system was initially generated via system builder by placing the complex inside an orthorhombic simulation cell containing solvent (TIP3P water model) and neutralized with counter-ions in addition to 0.15 M NaCl to represent physiological conditions. The system was parameterized with OPLS_2005 force field. The first step was an energy minimization in the steepest descent direction to relieve the steric clash, followed by a step-by-step equilibration of the system using NVT and NPT ensembles. Production MD was performed for 500 ns at 300 K and 1 atm using the Nose–Hoover thermostat and Martyna–Tobias–Klein barostat, with periodic boundary conditions. The trajectories were saved every 100 ps for post-analysis – RMSD, RHF, Rg (radius of gyration), formation of hydrogen bonds, total energy and binding stability were analyzed – in order to obtain a comprehensive description of conformational dynamics and structural integrity of the protein–protein complex under near-physiological conditions [21].

Codon Optimization and In-Silico Cloning

The terminal amino acid sequence of the vaccine construct underwent reverse-translation into a nucleotide sequence utilizing EMBOSS Backtranseq (https://www.ebi.ac.uk/jdispatcher/st/emboss_backtranseq). Subsequently, the codon usage was optimized for expression in *E. coli* via either the ExpOptimizer tool ([Codon Optimization \(ExpOptimizer\) - Online Tools](#)). For the purpose of in-silico cloning, the optimized nucleotide sequence was introduced into plasmid vectors such as PET28a(+) using SnapGene software, thereby facilitating prospective experimental expression [22].

Immune Simulation

The C-ImmSim server (<https://kraken.iac.rm.cnr.it/C-IMMSIM/index.php>) was employed to simulate the immune response elicited by the vaccine construct. This computational tool models the activation of various immune cells, encompassing CD4+ helper T-cells, CD8+ cytotoxic T-cells, B-cells, as well as the production of immunoglobulins and cytokines such as IFN- γ and IL-2. The simulation juxtaposed responses with and without the adjuvant to assess its influence on immune activation and memory formation. These predictions offered insights into the vaccine's potential to provoke a robust and enduring immune response [5].

RESULTS

Retrieval and Selection of Target Proteins

The findings showed homo sapiens Glypican-3 (UniProt ID: P51654) (Taxonomy ID: 9606). This protein is tested on the protein level (PE = 1) and belongs to the version of the sequence SV = 1. The 3d structure is shown in **Figure 1**.

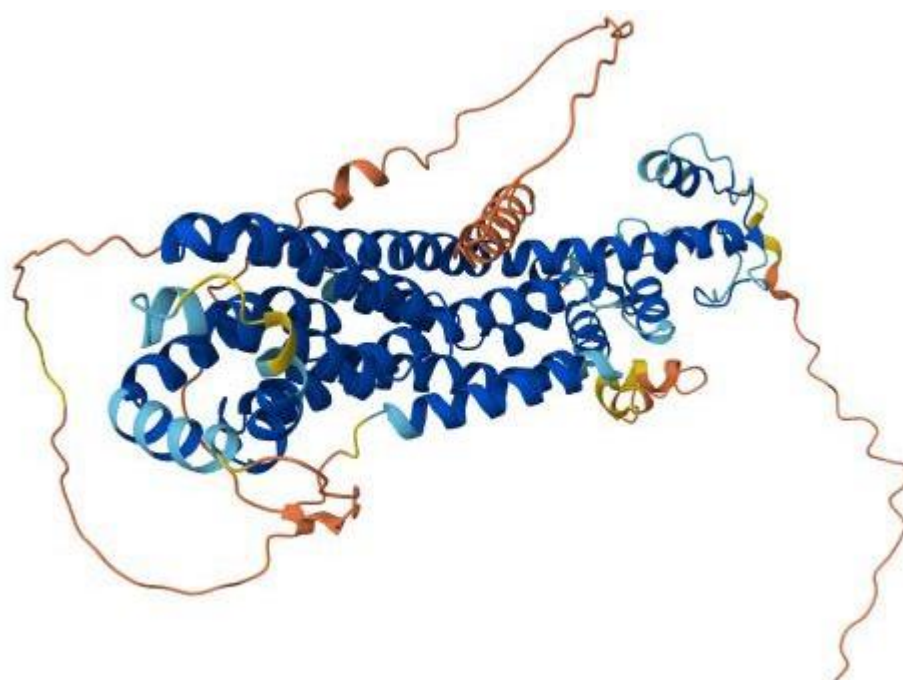


Figure 1: 3D structure of targeted protein.

Selection of targeted protein

Protein obtained a VaxiJen antigenicity score of **0.4495**, which makes it a probable antigen and thus has a strong indication of being an immunogen. It was proved to be non-allergenic and non-toxic, which is the mandatory safety prerequisite towards its development as a vaccine. TMHMM analysis estimated that the selected protein had no transmembrane helices. All residues (1 580) were confined to the extracellular region, meaning full accessibility to the surface and justifying its appropriateness in epitope targeting.

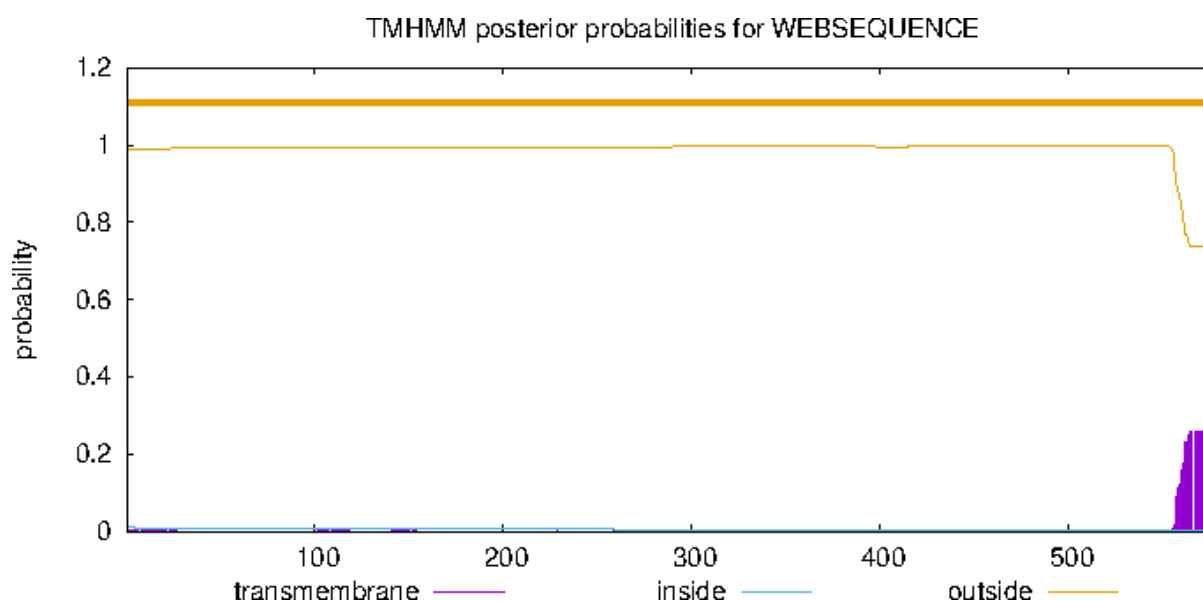


Fig 2: Subcellular location analysis of targeted protein.

Prediction of B-cell Epitopes

IEDB server was used to predict B-cell epitopes. The identified linear epitopes scored high in prediction and this proves that there is likelihood of recognition by the B-cell and production of antibodies that became more convincing in the immunogenicity profile of the construct Table 1.

Table 1: Selected B cell Epitopes

No.	Start	End	Peptide	Length
19	413	458	SPVAENDTLCWNGQELVERYSQKAARNGMKNQFNLHELKMKGPEPV	46

Prediction of MHC Class I Epitopes

Two epitopes generated by MHC class I were selected with reference to the binding affinity (IC50 value), percentile rank as well as immunogenicity. All of the epitopes were confirmed by VaxiJen, AllerTOP and ToxinPred and all of them stated that they were antigenic and non-toxic. These are the epitopes, which were anticipated to carry a high affinity to various HLA alleles, as illustrated in Table 2.

Table 2: Selected MHC Class I Epitopes

HLA Allele	Protein No.	Start	End	Epitope	Length	IC50 Score
HLA-A*02:01 HLA-A*68:02 HLA-A*02:03 HLA-A*02:01	3	49	57	ELFDSLFPV	9	3.02
HLA-A*02:01 HLA-A*68:02 HLA-A*02:03 HLA-A*02:01	3	49	58	ELFDSLFPVI	10	28.41
HLA-A*31:01	9	32	41	GMIKVKNQLR	10	70

Prediction of MHC Class II Epitopes

The choice of the two MHC class II epitopes was identified to have undergone adequate binding recognition of broad range of alleles, through the assistance of anti-genicity and low values of IC50. Epitopes were screened in the allergenicity and toxicity filters and were suitable to be used in designing the vaccine. Table 3 includes the list of the chosen epitopes.

Table 3: Selected MHC Class II Epitopes

HLA Allele	Protein No.	Start	End	Epitope	Length	IC50 Score
HLA-DPA1*01:03/DPB1*02:01 HLA-DPA1*03:01/DPB1*04:02 HLA-DPA1*02:01/DPB1*01:01	3	43	57	ELFDSLFPV	15	31.2
HLA-DRB1*01:01 HLA-DRB1*11:01	10	9	23	FHNLGNVHS	15	5.8
HLA-DRB1*04:05	6	19	33	FSTIHDSIQ	15	49.5

Vaccine Construct Assembly

Specific linkers were utilized to assemble the designed vaccine construct in order to provide structural stability and correct folding. MHC class I and II epitopes were fused in SSL and AAY linkers which give flexibility and optimum separation of space respectively. Immunostimulatory adjuvant Patent US11077184 was introduced at the N-terminus and C-terminus with EAAK linker, which improves immune stimulation. A 6xHis tag was inserted at the C-terminus using KK linker along with a cleavage/removal signal (VS) to improve the downstream expression and purification. The general design reduced steric hindrance, preserved epitope accessibility and retention of structural integrity of the vaccine constructs as shown in fig 3.

MEAAAKFNNFTVSVFRLVPKVSASHLEAAYSPVAENDTLCWNGQELVERYSQKAARNGMKNQFNLHELKMKGPEPVSSLELFDSLFPVSSLGMIKVKNQLRSSLNELFDSLFPVGPGPELFDSLFPVGPGPFHNLGNVHSGPGPGFSTIHDSIQEAAAKFNNFTVSVFRLVPKVSASHLEKKHHHHHH

Fig 3. Schematic representation of the multi-epitope vaccine construct. Blue represent B cell epitope, Gray represent adjuvant; Yellow segments represent linker sequences. Green regions indicate MHC class I epitopes. Pink regions represent MHC class II epitopes, and the dark green segment corresponds to the 6x-His tag

Vaccine Allergenicity, Antigenicity, and Solubility Prediction

The finalized vaccine construct was predicted to be non-allergenic and non-toxic, as determined by AllerTOP and ToxinPred,

respectively. Antigenicity analysis using VaxiJen yielded a score of **0.6422**, indicating strong immunogenic potential. Additionally, Protein-Sol predicted a scaled solubility score of 0.492, suggesting favorable solubility under physiological conditions as shown in **figure 4**. The theoretical isoelectric point (pI) was calculated to be 8.290, supporting its suitability for expression and formulation in vaccine development.

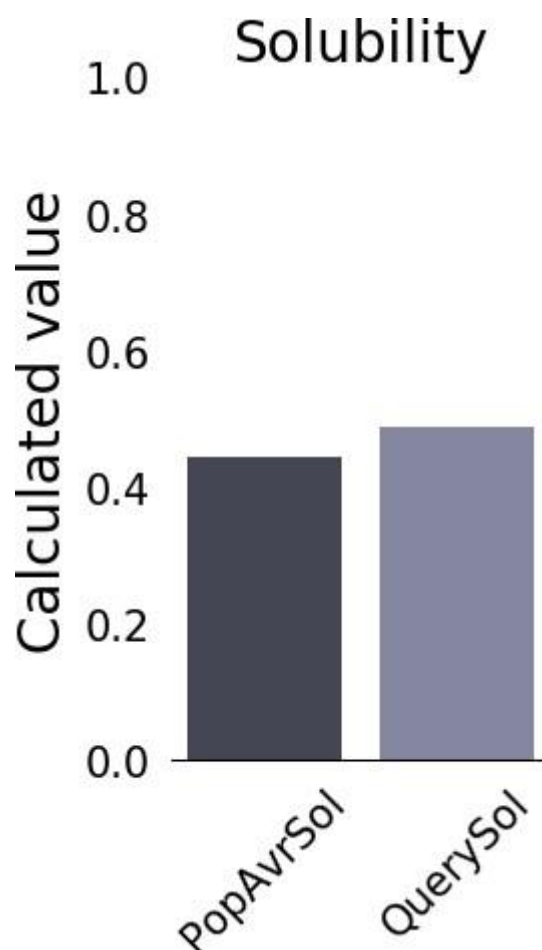


Fig 4. solubility level

Physicochemical Properties

The resulting vaccine construct was an almost neutral, 190 amino acid vaccine construct with a molecular weight of 21,057.85 Da, and a theoretical pI of 7.23. The protein sequence had the same amount of negatively charged (Asp + Glu = 17) and positively charged residues (Arg + Lys = 17), implying a proportionate distribution of charges. Its atomic structure was estimated as C955H1442N262O269S5 which equates to a total of 2,933 atoms. Its extinction coefficient was determined to be 19,480 M⁻¹cm⁻¹, and absorbance value at the concentration of 0.1 per cent (1 g/L) in water under oxidized and reduced conditions of cysteine to be 0.925.

The instability index (39.11) indicated the construct was stable whereas aliphatic index (73.37) indicated good thermostability. GRAVY score = -0.328 indicated a generally hydrophilic nature, which increased its solubility. It was approximately 30 h in mammalian reticulocytes (in vitro), >20 h in yeast (in vivo), and >10 h in *Escherichia coli* (in vivo), which allows inferring that the construct can be expressed successfully in several hosts.

Tertiary Structure Prediction and Validation

The **tertiary structure of the vaccine construct** was successfully predicted, exhibiting a compact and well-ordered folding pattern with appropriate spatial organization of structural domains, epitopes, and linker regions. Structural validation using the **SAVES v6.0 server** confirmed the stereochemical quality of the model. **Ramachandran plot analysis** showed that **81.2%** in favored region, 18.1% allowed region, 0.6% generously and 0% disallowed region, Allowed region. Importantly, no residues were observed in generously allowed or disallowed regions, confirming the high stereochemical accuracy and backbone geometry of the modeled construct.

Further validation demonstrated a **Z-score of -2.19**, which falls within the range of experimentally determined native proteins of similar size, supporting the model's reliability. Additionally, **ERRAT analysis yielded a quality factor of 91.86**, further confirming the structural integrity and accuracy of the predicted vaccine model.

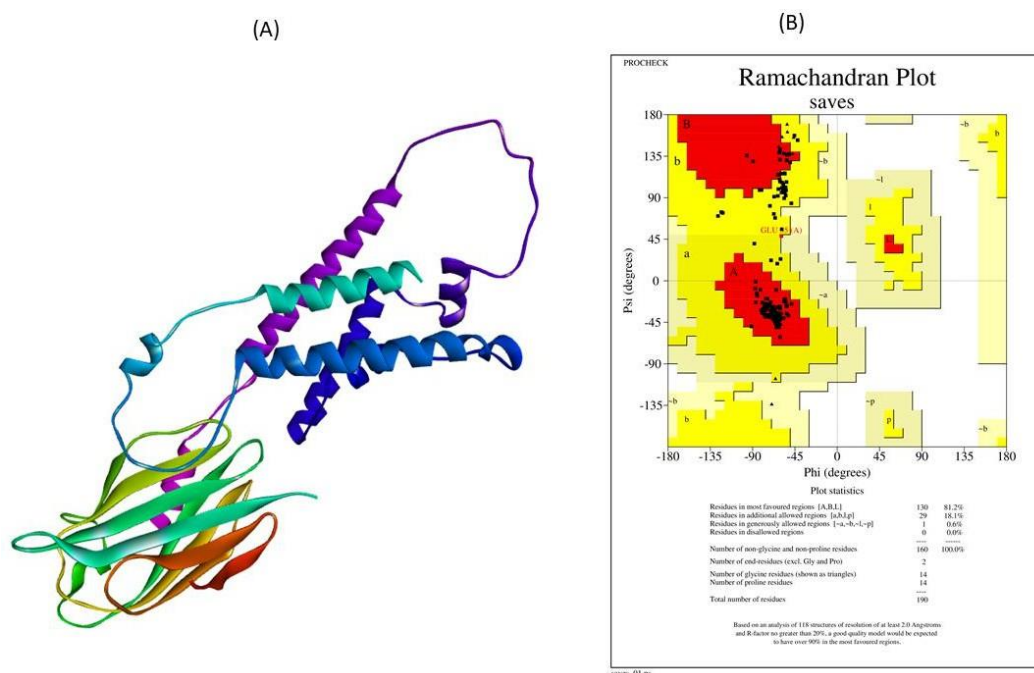


Fig 6. (A), 3D structure prediction. (B), Ramachandran Plot

Immune Interaction Analysis

ClusPro 2.0 molecular docking of the vaccine construct with human TLR4 produced a few stable complexes. Cluster 0 had the highest reliability and the highest number of members (429 total members) and the highest weighted score (-1228.1), whereas Cluster 1 (-1043.3) and Cluster 2 (-1129.2) both displayed strong binding energies and indicated that consistent interaction occurred across conformations.

The subsequent LigPlot+ analysis showed that there are numerous stabilizing interaction points such as all key hydrogen bonds (Lys97, Lys95, Arg92, Tyr105, Ser110, Glu200) and the extensive hydrophobic contacts (Val167, Leu169, His69, Ile83, Tyr108). All these interactions suggest a high-affinity binding geometry and structural complementarity between the vaccine construct and TLR4 indicating its potential to effectively stimulate immune signaling as shown in Figure 7.

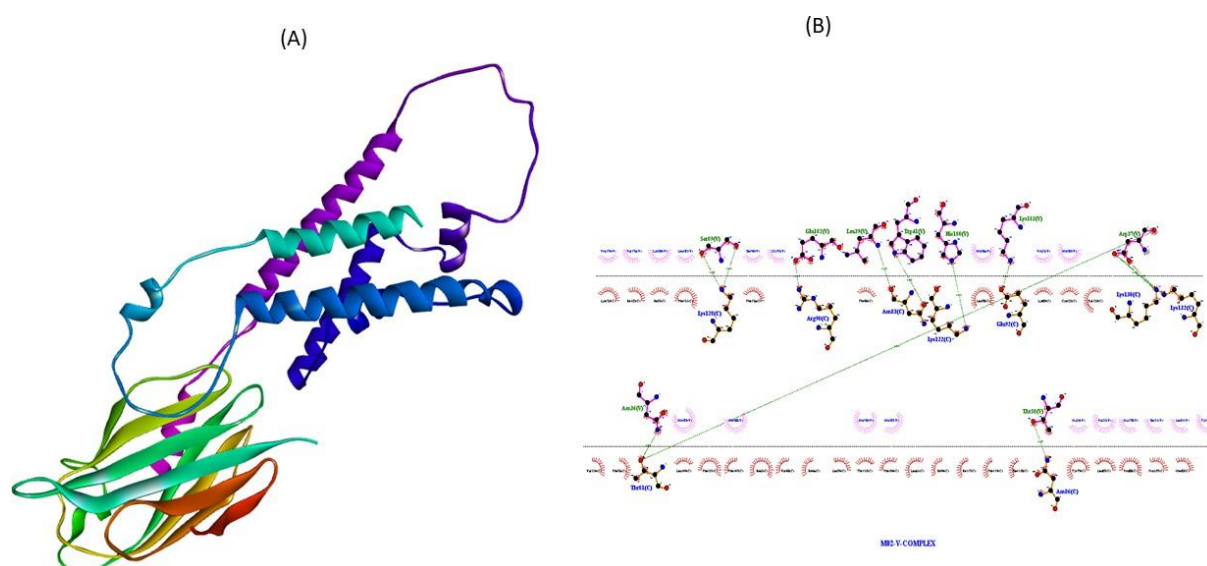


Fig 7. (A), Docking complex and (B), interaction analysis

Molecular Dynamic (MD) Simulations

The receptor fluctuates the Rg within a narrow range of $\sim 15.8\text{--}18.4\text{ \AA}$ with a long-run average around $\sim 16.8\text{ \AA}$ and without any directional drift (Fig 2), suggesting that RMSDs are accurate predictions of a well-packed compact fold that neither expands nor contracts along the targeted simulations. Except for an initial very short and gently contracting period which lasts until $\sim 20\text{--}30\text{ ns}$, the profile reaches a steady state of fluctuations around its mean through to the end of the simulation. This type of behavior is characteristic of a native-like core with loop breathing being the only source of fluctuation; and crucially as there's no gradual further compaction or swelling that would be indicative of misfolding or partial unfolding on binding it does not appear to significantly destabilize the receptor scaffold.

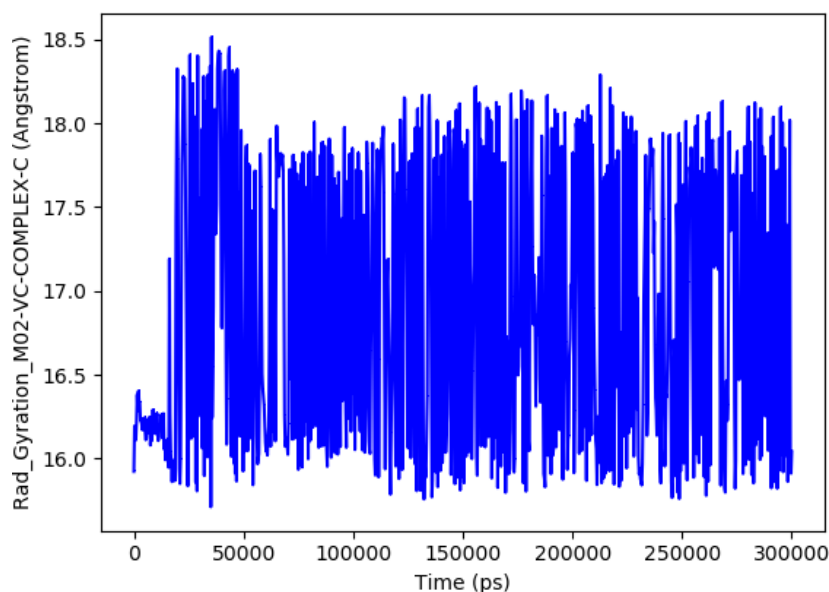


Fig 8: Radius of gyration of C chain.

On the other hand, chain V presents an initially expanded and extended state ($R_g \sim 22.5\text{--}23.0\text{ \AA}$) that progressively tightens to $\sim 19.5\text{--}20.5\text{ \AA}$ by $\sim 120\text{--}150\text{ ns}$ with a subsequent leveling off albeit on top of moderate fluctuations and occasional brief deviations around $\sim 21\text{ \AA}$, suggestive of an induced ordering upon binding mechanism: loose ensemble remodels, removes additional solvent exposure and adopts a tighter conformation optimized for the interface engagement. Finally, lack of reopening indicates that upon establishment of major interfacial contacts the peptide construct becomes non-exploring rather than exploring open states.

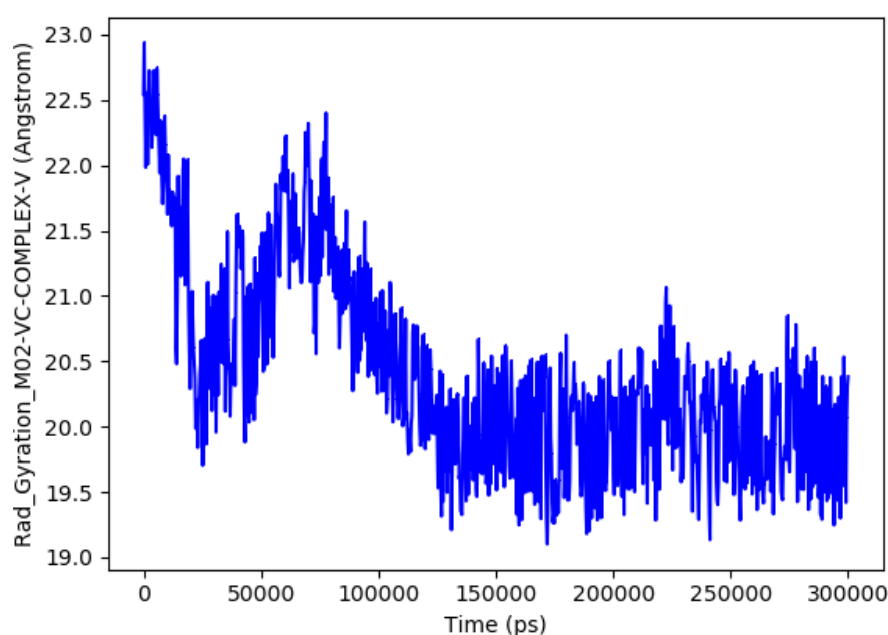


Fig 9: Radius of gyration V chain.

RMSD depicts rapid initial relaxation and sustained plateaus for binding strains of both partners. This is in agreement with moderate domain/loop rearrangements while maintaining the global structure for both chains. Chain V exhibits a greater initial rearrangement (jumping on to ~8.5–9.2 Å within the first few ns), followed by stability in that region, consistent with compaction observed from Rg. The combination—small, steady C-side RMSD and transient followed by stable V-side rearrangement—supports an induced-fit model wherein the vaccine chain undergoes a binding-interface reorganization with the receptor backbone array largely preserved.

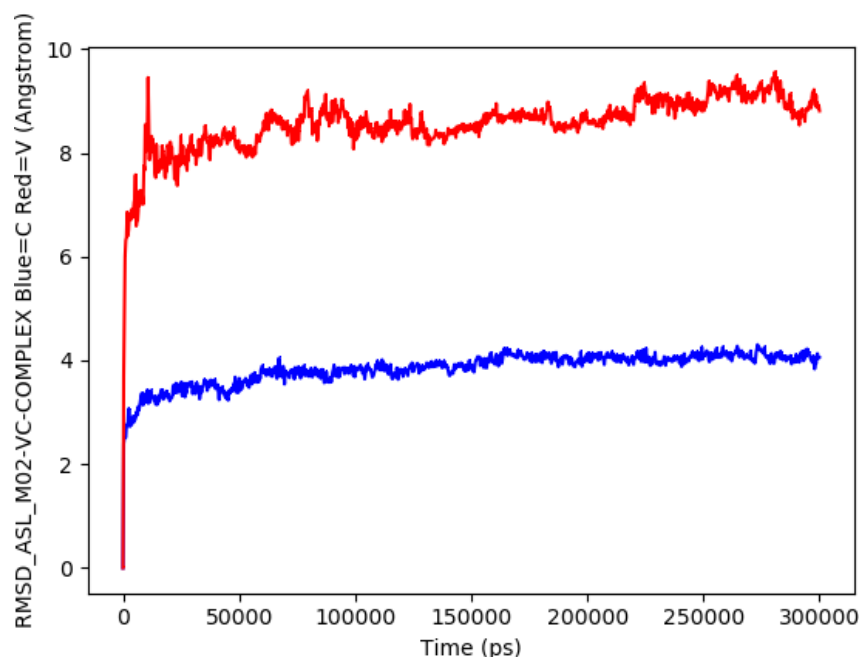


Fig 10: RMSD – Blue = Chain C (receptor), Red = Chain V (vaccine).

By and large, the per-residue fluctuations for the receptor are low (in the ~1.0–2.0 Å range) outside a few loop/terminus spikes—there's ones around residues ~30–40, ~110–135, and again at the C-terminal tail (~160) that reach up to 3~4 Å. That speaks of a rigid core with some mobility in loops nestled away from solvent—typical fare for a stable globular protein bound to something. There is no widespread increase in RMSF over all secondary structure elements, which indicates that the binding event does not carry destabilizing or spreading motions through the receptor helices/strands; flexibility continues to be concentrated to expected but non-catastrophic regions.

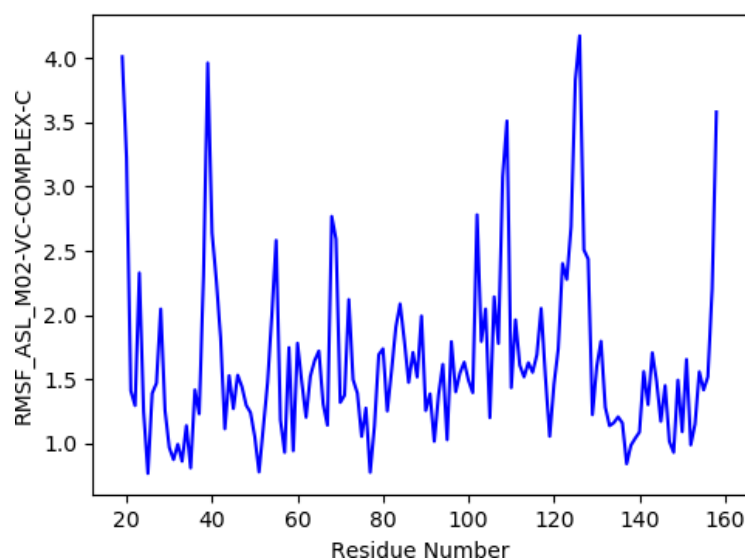


Fig 11: RMSF – Chain C (per residue fluctuation).

The vaccine/peptide construct has greater and more diverse mobility with the prominent peaks at ~30–50 Å, ~120–130 Å (~7~8 Å) and a moderate level of flexibility between ~150–165 (~4~5 Å), overlying a baseline of 1.5 to 3 Å over the rest of the protein. These hotspots often coincide with linker/loop regions and solvent exposed epitope regions, can aid in the early phase of interface

searching, and can continue as partially dynamical even after binding. The lack of extensive high RMSF throughout the entire chain indicates that only certain regions maintain flexibility and other regions are probably locked into position with receptor contacts.

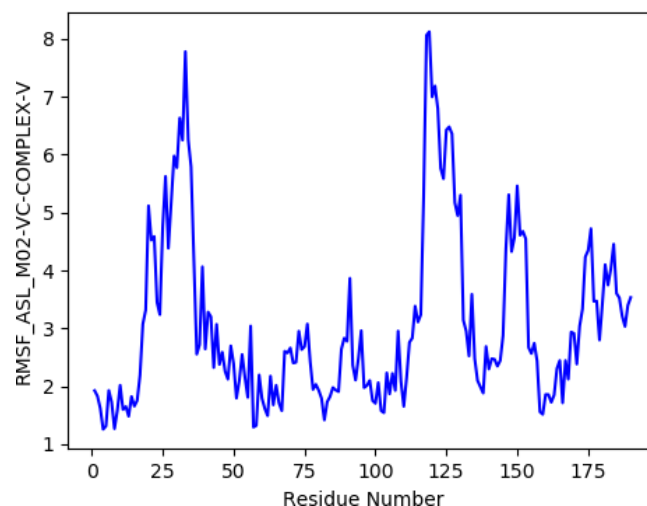


Fig 12: RMSF - Chain V (residual mobility).

The total solute energy drops from approximately $-200 \text{ kcal}\cdot\text{mol}^{-1}$ at the beginning to $\sim -900 \text{ kcal}\cdot\text{mol}^{-1}$ around $\sim 150\text{--}200 \text{ ns}$, and fluctuates in that lower “favorable” band for the next 300 ns. While the absolute values will depend on the force field and what is included in the “solute” term, a monotonic descent followed by a well with low energy is consistent with progressive improvement of packing/solvation and development of inter-molecular contacts. Most importantly, there is no long-term upward drift to suggest destabilization or incipient dissociation; instead the system simply evolves into a steady-state energetic minimum.

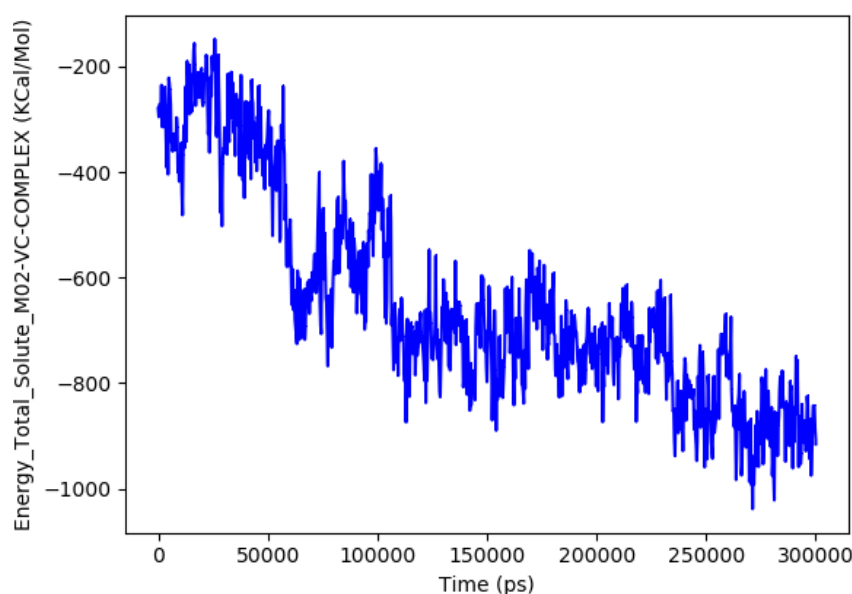


Fig 13: The overall solute energy of the complex.

The H-bond count regularly grows from an initial $\sim 8\text{--}12$ to a more crowded $\sim 12\text{--}18$ band at about $100\text{--}150$ or so ns, with spikes of up to $\sim 20\text{--}21$ at some points near the end. This growing and then constant H-bond network correlates with the compaction of chain V and the plateauing RMSDs, reflecting the establishment and maintenance of selective polar contacts at the interface. The greater number of H-bonds and reduced total solute energy is consistent with the complex “locking in” over time as transient early contacts are exchanged for a more radiative, dense network of interactions that drive the observed structural stability.

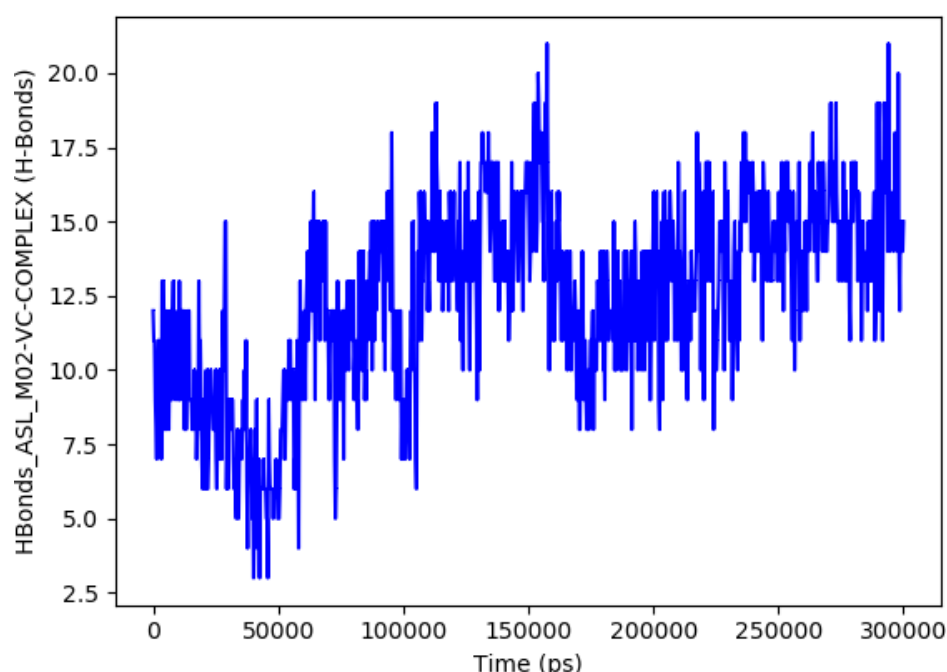


Fig 14: Intermolecular hydrogen bonding in the complex.

Codon Optimization and In-Silico Cloning

Codon optimization considerably enhanced the expression efficiency of the designed vaccine construct in *E. coli*. The Codon Adaptation Index (CAI) rose to 0.87, indicating that it is highly compatible with the host translational machinery. By analogy, the GC content was decreased to 53.86% to ensure that it fell within the optimal range of bacterial expression and minimises the chance of mRNA instability. The gene was optimized and predicted to produce a protein of 190 amino acid with a molecular weight of 21,042.3 Da, which is believed to be good to express and retain in a prokaryotic system. The construct was then adequately optimized and it was inserted successfully into pET28a(+) vectors thus indicating its preparedness of to be subjected to downstream experimental validation as shown in **Fig 15**.

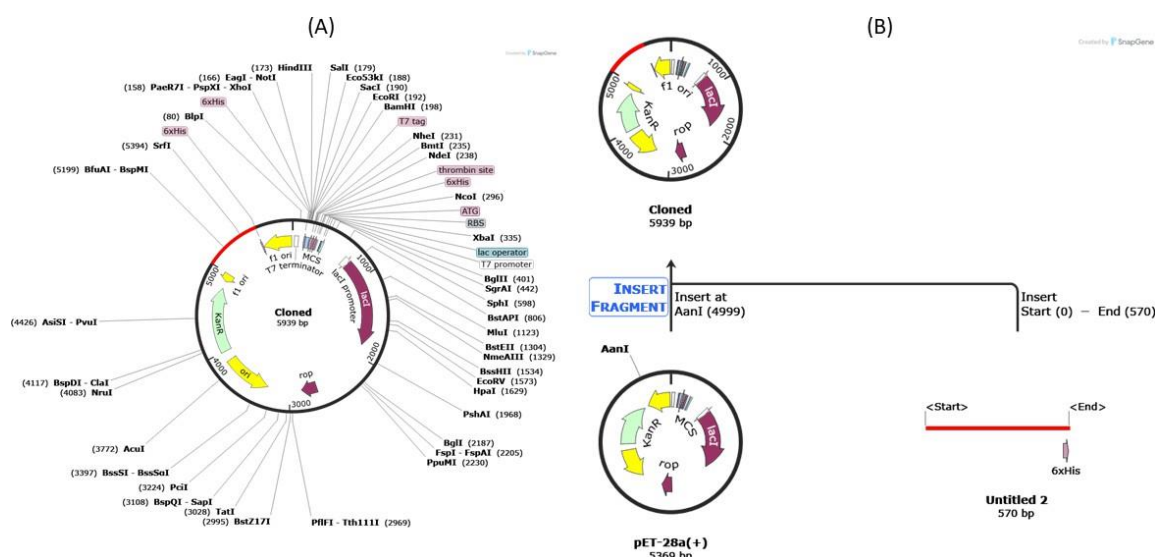


Fig 15: Vector cloning of vaccine construct. (A) Cloned vector illustration. (B) History of cloned vector.

Immune Simulation

This graph shows the peak of lymphocytes vi.a B-cell vi.b helper T cell (CD4+) and xiii.T cell lytic (of CD8+) increase after the

vaccination. The hyperbolic growth of helper T-cells whole results in growth of the B cells in proliferation to explain the concerted action of the body to sustain this high level of antibody production as explained in Fig 8. The outcome of the boosted cytotoxic T cell confirms that the vaccine does actually have the capability to stimulate cell mediated immunity required not just in the attitude of destroying the infected host cells, but in the feeling of protection in general. In the humoral versus cellular response pendulum response occurs as one way capacity that states the entire production of the general immune induction by the vaccine, and this is what an ideal vaccine film involves.

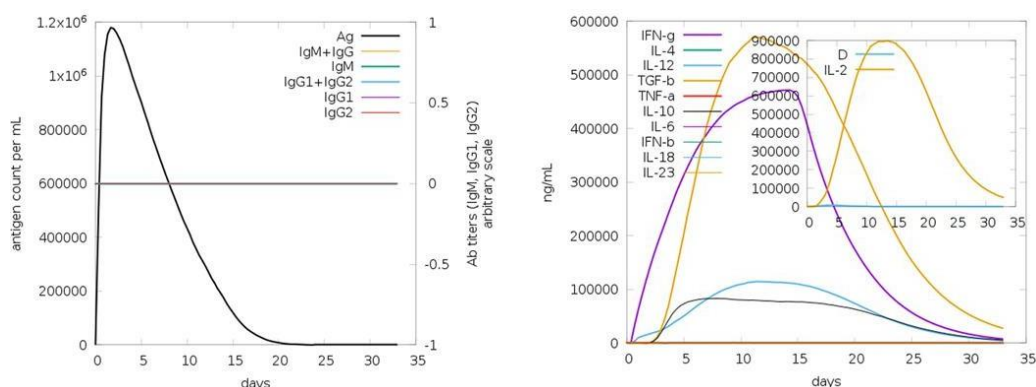


Fig 16: B-cell and T-cell Population Expansion Post-immunization

We have herein demonstrated in this Fig 9, that, the element of time enters into the exposure of the antigen, and into the formation of the antibodies which are formed at the time of the stimulation of the special vaccine with which we deal. The antigen presentation comes first and is both directional and immediate with regards to immunization and the sharp reduction is due to the destroying of the antigen by the immune system. Meanwhile, the titers of the antibodies are soaring to great heights and justify an adequate proliferation of the B-cell and the production of the plasma cell. According to kinetics, the vaccine might be endowed with the affable agency of arousing humoral immunity; the reaction to the vaccine and the neutralization of the antigen in the body, inexhaustibly stalemate memories baseline. The fact that the long term produced long lasting titers of the ant antibodies despite the peak of the initial stage is lost shows probability that the protective immunity is long term.

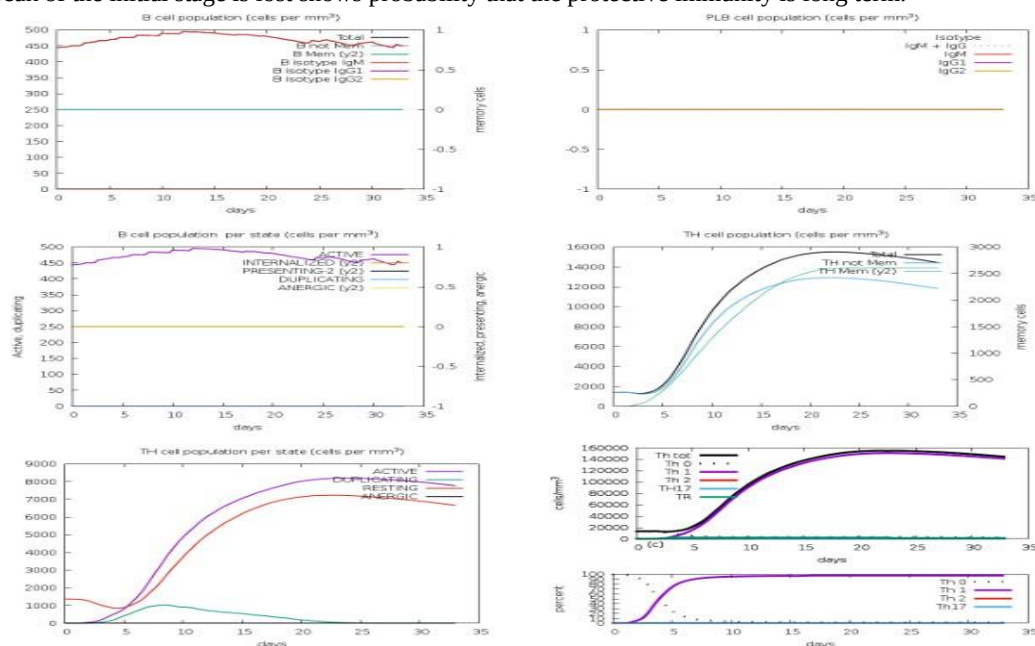


Fig 17: Antigen and Antibody Concentration Dynamics Following Vaccination.

Fig 18 illustrates modification of time of cytokine release and generation of memory cells following immunization. It would only mean positive rate of activation of the T-cells when positive changes to the IL-2 and IFN- γ cytokines levels are observed and positive changes of the IL-4 and IL-10 would indicate a successful rate of modulation and stimulation of the B-cells differentiation. It is also intriguing to mention that the cytokine concentration is vindicated in the later stages of the simulation, and T- and B-membranes are generated. One of the implications of this is that the immune system shifts to the memory phase rather than the effector phases that play a significant role in long term immunity. The existence of the longlived memory cells would suggest that the immunity system will indicate the recollection to the introduction to the pathogen with the high rate and intensity in an incident that might show the re-encounter.

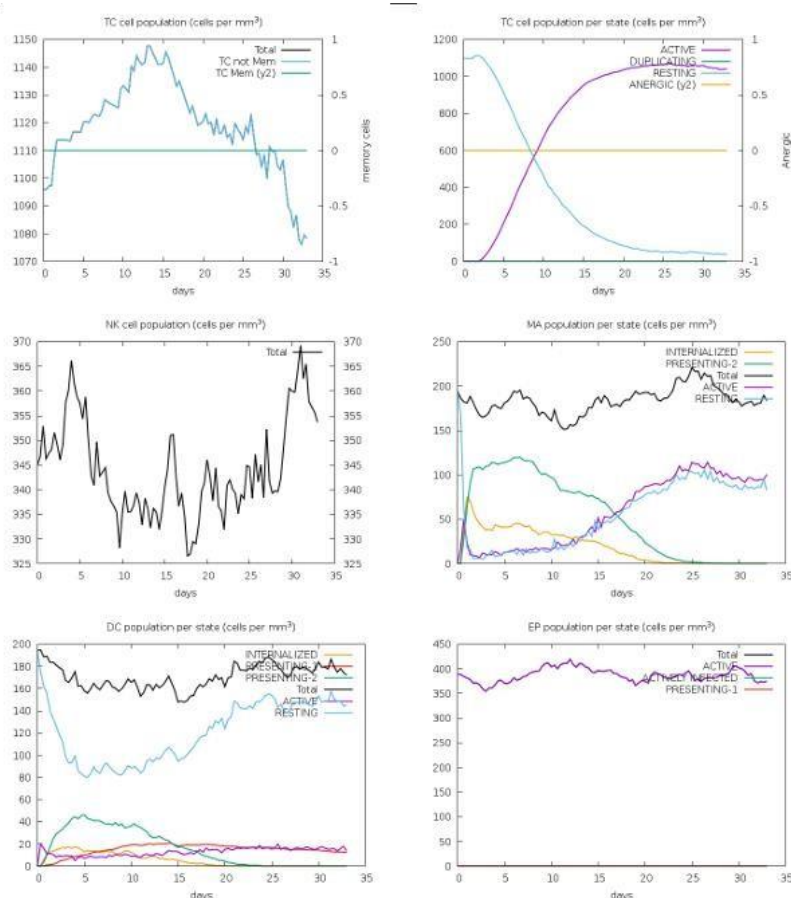


Fig 18: Cytokine Profile and Immune Memory Establishment.

DISCUSSION

HCC is still ranked among the deadliest tumors, because of late detection, high recurrence rate and therapy resistance. Standard options have failed to extend overall survival significantly beyond stage-adjusted nuanced treatments consistent with the urgency for new novel immunotherapeutic approaches [23]. In this regard, glypican-3 (GPC3) is increasingly recognized as a promising tumor antigen for its limited expression in normal adult tissues and high levels of over-expression in HCC. Our investigation exploited GPC3 as the anchor in developing a multi-epitope vaccine, targeted at engaging the humoral and cellular aspects of immunity, which potentially offers stronger and more sustainable protection than current therapeutic approaches [24].

Epitopes were chosen according to their high antigenicity, non-allergenicity and non-toxicity, thus guaranteeing the safety as well as immunogenicity of the construct. B-cell epitopes were designed to activate potent antibody responses, whereas CTL and HTL epitopes appealed to the killing T cells, helper T cell activation, respectively. Structural stability and immunostimulatory activity were augmented by the addition of flexible linkers, adjuvants, and purification tags. Crucially, the rationally designed construct possessed a high VaxiJen score (0.6422), validating its immunogenicity potential and retained potent physicochemical attributes such as hydrophilicity, thermostability, and solubility. These results suggest that the vaccine has a strong potential of being efficiently expressed, purified and tested in subsequent experimental models [25].

Stereochemical correctness and structural stability of the vaccine were further substantiated by structural modeling and validation. Specifically, the Ramachandran plot analysis identified over 81% of residues in favored regions, ERRAT quality factor exceeded 90, and Z-score was within the experimentally validated proteins standard range. Such structural parameters suggest that the designed vaccine is stereo chemically accurate and structurally robust and thus is highly likely to fold properly and present the epitopes in their native-like conformation [26]. As already mentioned, the vaccine conformation consistency is key to epitope accessibility, which allows for successful receptor recognition preconditions for a robust immune response. Importantly,

molecular docking and molecular dynamics simulations also verified that the vaccine is highly likely to succeed based on theoretical considerations [6].

First, the vaccine binding with TLR4 demonstrated binding energies, an abundance of hydrogen bonds, and hydrophobic interaction involving crucial residues as shown in Table 3. These attributes can be described as bound with strong affinity and structural integrity, which translated into successful completion of 500 ns molecular dynamics simulations with receptor's stability and vaccine chain compacting due to induced fit. Therefore, RMSD and Rg values indicated stable conformations, whereas RMSF uncovered highly flexible loop regions, which are to bind with the receptors without altering the overall immunogenic structural fold stability. Furthermore, the sequential reduction of total solute energy and increasing number of hydrogen bonds further indicated the final vaccine-receptor complex's thermodynamic and respective functional stability [22].

The codon optimization and In-silico cloning enhanced the translational potential of vaccine construct. The high value of CAI (0.87) along with optimal GC content (53.86%) supported the efficient expression of the wild-type heavy chain construct in *E. coli*, which is one of accepted desired hosts for production of recombinant proteins. Furthermore, immune simulations showed strong activation of B-cells, CD4⁺ helper T-cells and CD8⁺ cytotoxic T cells as well as long-lasting presence of antibodies and cytokines. The induction of long-lived memory cells implied that the vaccine has potential to provide lasting immune protection, which is a crucial attribute for successful cancer immunotherapy [6].

These findings suggest the promise in some rational vaccinology methodologies, such as the immunoinformatic driving tactic. The GPC3 made multi-epitope structure was highly immunogenic, had structural integrity, immune receptor binding stability and strong idealize immune simulations. When compared to the well-developed monoclonal antibody or CAR-T-based treatment, the computationally designed vaccine is also a cost-effective and optimal strategy to provide a generalized application independent of antigens with fewer off-target toxicity. In vivo and in vitro experiments are needed to prove these predictions, however this study offers a firm computational basis for the development of GPC3-targeted immunotherapy for HCC.

CONCLUSION

This work has also provided a detailed in silico design template for the development of an epitope-based multi-epitope vaccine against GPC3, which is a clinically relevant marker of hepatocellular carcinoma. The chimeric construct was shown to possess a high level of antigenicity and non-allergenicity, stable tertiary structure as well good stability according to several computational analysis including molecular docking, long-term dynamics simulations indicated that it had also favourable interactions with TLR4 and conformational stability. Codon optimization was conducted to make it compatible with the expression of recombinant protein and immune simulation revealed a balanced humoral–cellular immunity and long-lasting memory cell formation. Collectively, these findings indicate that this proposed vaccine could be a safe, cost-effective and immunogenic therapeutic strategy for the treatment of HCC. Although many challenges remain for experimental and clinical validation, the work herein demonstrates that immunoinformatic is a transformative tool to expedite cancer vaccine development and establishes a strong rationale for translating computational vaccine designs to precision immunotherapies.

REFERENCES

- [1] A. Ozakyol, "Global Epidemiology of Hepatocellular Carcinoma (HCC Epidemiology)," *J. Gastrointest. Cancer*, vol. 48, no. 3, pp. 238–240, Sept. 2017, doi: 10.1007/s12029-017-9959-0.
- [2] B. Douradinha, "Computational strategies in *Klebsiella pneumoniae* vaccine design: navigating the landscape of in silico insights," *Biotechnol. Adv.*, vol. 76, p. 108437, Nov. 2024, doi: 10.1016/j.biotechadv.2024.108437.
- [3] S. L. Ali *et al.*, "Promising vaccine models against astrovirus MLB2 using integrated vaccinomics and immunoinformatics approaches," *Mol. Syst. Des. Eng.*, vol. 9, no. 12, pp. 1285–1299, 2024, doi: 10.1039/D3ME00192J.
- [4] A. Devi and N. S. N. Chaitanya, "In silico designing of multi-epitope vaccine construct against human coronavirus infections," *J. Biomol. Struct. Dyn.*, vol. 39, no. 18, pp. 6903–6917, Dec. 2021, doi: 10.1080/07391102.2020.1804460.
- [5] S. Basak *et al.*, "In silico designing of vaccine candidate against *Clostridium difficile*," *Sci. Rep.*, vol. 11, no. 1, p. 14215, July 2021, doi: 10.1038/s41598-021-93305-6.
- [6] S. Arumugam and P. Varamballi, "In-silico design of envelope based multi-epitope vaccine candidate against Kyasanur forest disease virus," *Sci. Rep.*, vol. 11, no. 1, p. 17118, Aug. 2021, doi: 10.1038/s41598-021-94488-8.
- [7] M. W. Cho *et al.*, "Polyvalent Envelope Glycoprotein Vaccine Elicits a Broader Neutralizing Antibody Response but Is Unable To Provide Sterilizing Protection against Heterologous Simian/Human Immunodeficiency Virus Infection in Pigtailed Macaques," *J. Virol.*, vol. 75, no. 5, pp. 2224–2234, Mar. 2001, doi: 10.1128/JVI.75.5.2224-2234.2001.
- [8] The UniProt Consortium, "UniProt: the universal protein knowledgebase," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D158–D169, Jan. 2017, doi: 10.1093/nar/gkw1099.
- [9] I. Dimitrov, D. R. Flower, and I. Doytchinova, "AllerTOP - a server for in silico prediction of allergens," *BMC Bioinformatics*, vol. 14, no. S6, p. S4, Apr. 2013, doi: 10.1186/1471-2105-14-S6-S4.
- [10] I. A. Doytchinova and D. R. Flower, "VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines," *BMC Bioinformatics*, vol. 8, no. 1, p. 4, Dec. 2007, doi: 10.1186/1471-2105-8-4.
- [11] Y. Chen, P. Yu, J. Luo, and Y. Jiang, "Secreted protein prediction system combining CJ-SPHMM, TMHMM,

- p and PSORT,”
- Mamm. Genome*
- , vol. 14, no. 12, pp. 859–865, Dec. 2003, doi: 10.1007/s00335-003-2296-6.
- [12] S. K. Dhanda *et al.*, “IEDB-AR: immune epitope database—analysis resource in 2019,” *Nucleic Acids Res.*, vol. 47, no. W1, pp. W502–W506, July 2019, doi: 10.1093/nar/gkz452.
- [13] L. J. Abu-Raddad, H. Chemaitelly, and A. A. Butt, “Effectiveness of the BNT162b2 Covid-19 Vaccine against the B.1.1.7 and B.1.351 Variants,” *N. Engl. J. Med.*, vol. 385, no. 2, pp. 187–189, July 2021, doi: 10.1056/NEJMc2104974.
- [14] M. Hebditch, M. A. Carballo-Amador, S. Charonis, R. Curtis, and J. Warwicker, “Protein-Sol: a web tool for predicting protein solubility from sequence,” *Bioinformatics*, vol. 33, no. 19, pp. 3098–3100, Oct. 2017, doi: 10.1093/bioinformatics/btx345.
- [15] P. Artimo *et al.*, “ExPASy: SIB bioinformatics resource portal,” *Nucleic Acids Res.*, vol. 40, no. W1, pp. W597–W603, July 2012, doi: 10.1093/nar/gks400.
- [16] L. J. McGuffin, K. Bryson, and D. T. Jones, “The PSIPRED protein structure prediction server,” *Bioinformatics*, vol. 16, no. 4, pp. 404–405, Apr. 2000, doi: 10.1093/bioinformatics/16.4.404.
- [17] J. Jumper *et al.*, “Highly accurate protein structure prediction with AlphaFold,” *Nature*, 2021, doi: 10.1038/s41586-021-03819-2.
- [18] S. W. Park, B. H. Lee, S. H. Song, and M. K. Kim, “Revisiting the Ramachandran plot based on statistical analysis of static and dynamic characteristics of protein structures,” *J. Struct. Biol.*, vol. 215, no. 1, p. 107939, Mar. 2023, doi: 10.1016/j.jsb.2023.107939.
- [19] A. Alekseenko, M. Ignatov, G. Jones, M. Sabitova, and D. Kozakov, “Protein–Protein and Protein–Peptide Docking with ClusPro Server,” in *Protein Structure Prediction*, vol. 2165, D. Kihara, Ed., in Methods in Molecular Biology, vol. 2165., New York, NY: Springer US, 2020, pp. 157–174. doi: 10.1007/978-1-0716-0708-4_9.
- [20] A. C. Wallace, R. A. Laskowski, and J. M. Thornton, “LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions,” *Protein Eng. Des. Sel.*, vol. 8, no. 2, pp. 127–134, 1995, doi: 10.1093/protein/8.2.127.
- [21] M. Gillett *et al.*, “Delivering the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cost effectiveness analysis,” *BMJ*, vol. 341, no. aug20 1, pp. c4093–c4093, Aug. 2010, doi: 10.1136/bmj.c4093.
- [22] K. Abraham Peele, T. Srihansa, S. Krupanidhi, V. S. Ayyagari, and T. C. Venkateswarulu, “Design of multi-epitope vaccine candidate against SARS-CoV-2: a *in-silico* study,” *J. Biomol. Struct. Dyn.*, vol. 39, no. 10, pp. 3793–3801, July 2021, doi: 10.1080/07391102.2020.1770127.
- [23] Y. Wu, H. Liu, and H.-G. Ding, “GPC-3 in hepatocellular carcinoma: current perspectives,” *J. Hepatocell. Carcinoma*, vol. Volume 3, pp. 63–67, Nov. 2016, doi: 10.2147/JHC.S116513.
- [24] T. Nakatsura, K. Ofuji, K. Saito, and T. Yoshikawa, “Critical analysis of the potential of targeting GPC3 in hepatocellular carcinoma,” *J. Hepatocell. Carcinoma*, p. 35, May 2014, doi: 10.2147/JHC.S48517.
- [25] S. An *et al.*, “GPC3-targeted immunoPET imaging of hepatocellular carcinomas,” *Eur. J. Nucl. Med. Mol. Imaging*, vol. 49, no. 8, pp. 2682–2692, July 2022, doi: 10.1007/s00259-022-05723-x.
- [26] A. C. B. Antonelli *et al.*, “In silico construction of a multiepitope Zika virus vaccine using immunoinformatics tools,” *Sci. Rep.*, vol. 12, no. 1, p. 53, Jan. 2022, doi: 10.1038/s41598-021-03990-6.